R Commands [Linear Regression]

File: slr.csv and mtcars.csv

Sr No	Purpose	Command
1	Import file slr.csv	
2	Plot scatter diagram	`> plot(slr$Advt, slr$Sales,)`
3	Find correlation	`> cor(slr$Advt, slr$sales)` `[1] 0.9491666`
4	Build regression model and find summary	`> mod = lm(slr$Sales~slr$Advt)` `> summary(mod)`
5	Find predicted values	`> predict(mod)`

6	Insert predicted values in data file	`> pred = predict(mod)`
		`> slr$predicted = NA`
		`> slr$predicted = pred`

7	Find error terms	`> slr$error = mod$residuals`
8	Install package 'car' for finding Durbin Watson Statistics	`> library(car)` `> dwt(mod)`
9	Add best fit line in scatter plot	`> plot(slr$Advt, slr$Sales, abline(lm(slr$Sales~slr$Advt),` `col = "red"))`
10	Check assumption 1: Linearity	`> plot(slr$Advt, slr$error, xlab="Advertisement",` `ylab = "Residuals", main = "Linearity")`
11	Check assumption 1: Constant Error Variance	`> plot(slr$predicted, slr$error, xlab="Predicted", ylab` `= "Residuals", main = "Constant Error Variance")`
12	Check assumption 3: Independence of Error	`> plot(slr$observation.no, slr$error, xlab="Observation` `No", ylab = "Residuals", main = "Independence of Error")`
13	Check assumption 4: Normality	`> hist(slr$error, xlab = "Residuals", main = "Histogram` `of Residuals", col="yellow")`

14	Save predicted values and error terms in a new file slr1	<pre>> slr1<-slr
		> head(slr1)
		  Observation.no Advt Sales predicted error
		1 1 92 930 902.3965 27.60349
		2 2 94 900 940.5374 -40.53740
		3 3 97 1020 997.7487 22.25127
		4 4 98 990 1016.8192 -26.81917
		5 5 100 1100 1054.9601 45.03994
		6 6 102 1050 1093.1009 -43.10094
		> write.csv(slr1,"C:/Users/inurture1/Desktop/slr1.csv")</pre>

15	Multiple Regression
	Import mtcars.csv
	Install package 'car' then call 'car'
		<pre>> mtcars <- read.csv("C:/Users/inurture1/Desktop/mtcars.csv")
		> View(mtcars)
		> install.packages("car")
		> library(car)</pre>

16	Build multiple regression model	<pre>> fit<- lm(mpg~ disp+hp+wt+drat, data = mtcars)
		> fit</pre>

17	Find summary	<pre>> summary(fit)</pre>

18	Find variance inflation factor	<pre>> vif(fit)
		    disp hp wt drat
		8.209402 2.894373 5.096601 2.279547</pre>

19	Find vif>5	<pre>> vif(fit)>5
		 disp hp wt drat
		 TRUE FALSE TRUE FALSE</pre>